WHAT IS CLAIMED IS:

1. A method for determining whether a cellular constituent is causal for a trait of interest **T**, the trait of interest **T** exhibited by one or more organisms in a plurality of organisms of a species, the method comprising:

(A) identifying one or more loci in the genome of said species, wherein each locus Q of said one or more loci is a site of colocalization for (i) a respective abundance quantitative trait locus (eQTL) genetically linked to a variation in abundance levels of the cellular constituent across the plurality of organisms and (ii) a respective clinical quantitative trait locus (cQTL) that is genetically linked to a variation in said trait of interest **T** across said plurality of organisms; and

(B) testing, for each respective locus Q of said one or more loci, whether (i) a genetic variation of said respective locus Q across said plurality of organisms, and (ii) said genetic variation in said trait of interest **T** across said plurality of organisms are correlated conditional on said variation in abundance levels of the cellular constituent across said plurality of organisms,

wherein, when (i) the genetic variation of one or more Q tested in (B), and (ii) the variation in said trait of interest **T** across said plurality of organisms are uncorrelated conditional on said variation in abundance levels of the cellular constituent across said plurality of organisms, said cellular constituent is determined to be causal for said trait of interest **T**.

2. The method of claim 1, the method further comprising, prior to said identifying, a step of determining a respective eQTL at a locus Q of said one or more loci using a first quantitative trait locus (QTL) analysis, wherein said first QTL analysis uses a plurality of abundance statistics for said cellular constituent as a quantitative trait, and wherein each abundance statistic in said plurality of abundance statistics represents an abundance value for said cellular constituent in an organism in said plurality of organisms.

3. The method of claim 2, the method further comprising a step of determining a respective cQTL at a locus Q of said one or more loci using a second QTL analysis, wherein said second QTL analysis uses a plurality of phenotypic values as a quantitative trait, each phenotypic value in said plurality of phenotypic values corresponding to an organism in said plurality of organisms.

4. The method of claim 1, wherein said respective eQTL and said respective cQTL are deemed to be colocalized at a locus Q of said one or more loci when said respective eQTL and said respective cQTL are within 3 cM of the locus Q.

5. The method of claim 1, wherein said respective eQTL and said respective cQTL are deemed to be colocalized at a locus Q of said one or more loci when said respective eQTL and said respective cQTL are within 1 cM of the locus Q.

6. The method of claim 3, wherein said first QTL analysis and said second QTL analysis each use a genetic map that represents the genome of said species.

7. The method of claim 6, which further comprises, prior to said identifying, a step of constructing said genetic map from a set of genetic markers associated with said species.

8. The method of claim 7, wherein said set of genetic markers comprises single nucleotide polymorphisms (SNPs), microsatellite markers, restriction fragment length polymorphisms, short tandem repeats, DNA methylation markers, sequence length polymorphisms, random amplified polymorphic DNA, amplified fragment length polymorphisms, or simple sequence repeats.

9. The method of claim 7, wherein genotype data is used in said constructing and wherein said genotype data comprises knowledge of which alleles, for each marker in said set of genetic markers, are present in each organism in said plurality of organisms.

10. The method of claim 7, wherein said plurality of organisms represents a segregating population and pedigree data are used in said constructing step, and wherein said pedigree data show one or more relationships between organisms in said plurality of organisms.

11. The method of claim 10, wherein said plurality of organisms comprises an F_2 population, a F_r population, a $F_{2:3}$ population, or a Design III population and said one or more relationships between organisms in said plurality of organisms indicates which organisms in said plurality of organisms are members of said F_2 population, said F_r population, said $F_{2:3}$ population, or said Design III population.

12. The method of claim 1 wherein said plurality of organisms is derived from a predetermined set of individuals.

13. The method of claim 1 wherein said plurality of organisms is derived from a predetermined set of strains.

14. The method of claim 13 wherein said set of strains is between 2 strains and 100 strains.

15. The method of claim 13 wherein said set of strains is between 5 strains and 500 strains.

16. The method of claim 13 wherein said set of strains is more than five strains.

17. The method of claim 13 wherein said set of strains is less than 1000 strains.

18. The method of claim 13 wherein said set of strains is diverse with respect to a complex phenotype associated with human disease.

19. The method of claim 13 wherein said set of strains is between 2 strains and 10 strains that, collectively, are diverse with respect to a complex phenotype associated with a human disease.

20. The method of claim 19 wherein said human disease is obesity, diabetes, atherosclerosis, metabolic syndrome, depression, anxiety, osteoporosis, bone development, asthma, or chronic obstructive pulmonary disease.

21. The method of claim 1 wherein said plurality of organisms is derived from crossing a predetermined set of strains.

22. The method of claim 22 wherein said plurality of organisms is an F_2 intercross, a backcross, or an F_2 random mating.

23. The method of claim 1 wherein the plurality of organisms is more than 1,000 organisms.

24. The method of claim 1 wherein the plurality of organism is between 100 organisms and 100,000 organisms.

25. The method of claim 1 wherein the plurality of organisms is less than 500,000 organisms.

26. The method of claim 1 wherein the plurality of organisms is between 5,000 and 25,000 organisms.

27. The method of claim 2, wherein each said abundance value is a normalized abundance level measurement for said cellular constituent in an organism in said plurality of organisms.

28. The method of claim 27, wherein each said abundance level measurement is determined by measuring an amount of said cellular constituent in one or more cells from said organism.

29. The method of claim 28, wherein said amount of said cellular constituent comprises an abundance of an RNA present in said one or more cells of said organism.

30. The method of claim 29, wherein said abundance of said RNA is measured by contacting a gene transcript array with said RNA from said one or more cells of said organism, or with nucleic acid derived from said RNA, wherein said gene transcript array comprises a positionally addressable surface with attached nucleic acids or nucleic acid mimics, wherein said nucleic acids or nucleic acid mimics are capable of hybridizing with said RNA species, or with nucleic acid derived from said RNA species.

31. The method of claim 27, wherein said normalized abundance level measurement is obtained by a normalization technique selected from the group consisting of Z-score of intensity, median intensity, log median intensity, Z-score standard deviation log of intensity, Z-score mean absolute deviation of log intensity, calibration DNA gene set, user normalization gene set, ratio median intensity correction, and intensity background correction.

32. The method of claim 2, wherein said first QTL analysis comprises:

(i) testing for linkage between (a) the genotype of said plurality of organisms at a position in the genome of said species and (b) said plurality of abundance statistics for said cellular constituent;

(ii) advancing the position in said genome by an amount; and

(iii) repeating steps (i) and (ii) until all or a portion of the genome of said species has been tested.

33. The method of claim 32, wherein said amount is less than 100 centiMorgans.

34. The method of claim 32, wherein said amount is less than 5 centiMorgans.

35. The method of claim 32, wherein said testing comprises performing linkage analysis or association analysis.

36. The method of claim 35, wherein said linkage analysis or association analysis generates a statistical score for said position in the genome of said species.

37. The method of claim 36, wherein said testing is linkage analysis and said statistical score is a logarithm of the odds (lod) score.

38. The method of claim 37, wherein said respective eQTL is represented by a lod score that is greater than 2.0.

39. The method of claim 37, wherein said respective eQTL is represented by a lod score that is greater than 4.0.

40. The method of claim 3, wherein said second QTL analysis comprises:

(i) testing for linkage between (a) the genotype of said plurality of organisms at a position in the genome of said species and (b) said plurality of phenotypic values;

(ii) advancing the position in said genome by an amount; and

(iii) repeating steps (i) and (ii) until all or a portion of the genome of said species has been tested.

41. The method of claim 40, wherein said amount is less than 100 centiMorgans.

42. The method of claim 40, wherein said amount is less than 5 centiMorgans.

43. The method of claim 40, wherein said testing comprises performing linkage analysis or association analysis.

44. The method of claim 43, wherein said linkage analysis or association analysis generates a statistical score for said position in the genome of said species.

45. The method of claim 44, wherein said testing is linkage analysis and said statistical score is a logarithm of the odds (lod) score.

46. The method of claim 45, wherein said respective cQTL is represented by a lod score that is greater than 2.0.

47. The method of claim 45, wherein said respective cQTL is represented by a lod score that is greater than 4.0.

48. The method of claim 1, wherein said plurality of organisms is human.

49. The method of claim 1, wherein said trait of interest **T** is a complex trait.

50. The method of claim 49, wherein said complex trait is characterized by an allele that exhibits incomplete penetrance in said species.

51. The method of claim 49, wherein said complex trait is a disease that is contracted by an organism in said plurality of organisms, and wherein said organism inherits no predisposing allele to said disease.

52. The method of claim 49, wherein said complex trait arises when any of a plurality of different genes in the genome of said species are mutated.

53. The method of claim 49, wherein said complex trait requires the simultaneous presence of mutations in a plurality of genes in the genome of said species.

54. The method of claim 49, wherein said complex trait is associated with a high frequency of disease-causing alleles in said species.

55. The method of claim 49, wherein said complex trait is a phenotype that does not exhibit Mendelian recessive or dominant inheritance attributable to a single gene locus.

56. The method of claim 49, wherein said complex trait is asthma, ataxia telangiectasia, bipolar disorder, cancer, common late-onset Alzheimer's disease, diabetes, heart disease, hereditary early-onset Alzheimer's disease, hereditary nonpolyposis colon cancer, hypertension, infection, maturity-onset diabetes of the young, mellitus, migraine, nonalcoholic fatty liver, nonalcoholic steatohepatitis, non-insulin-dependent diabetes mellitus, obesity, polycystic kidney disease, psoriases, schizophrenia, or xeroderma pigmentosum.

57. The method of claim 1, wherein said respective eQTL and said respective cQTL are deemed to be colocalized at a locus *Q* of said one or more loci when said respective eQTL and said respective cQTL are within 40 cM of the locus *Q*.

58. The method of claim 1, wherein said respective eQTL and said respective cQTL are deemed to be colocalized at a locus *Q* of said one or more loci when said respective eQTL and said respective cQTL are within 10 cM of the locus *Q*.

59. The method of claim 2 wherein said abundance value comprises an amount of said cellular constituent in a tissue of said organism, a concentration of said cellular constituent in a tissue of said organism, a cellular constituent activity level for said cellular constituent in a tissue of said organism, or the state of modification of said cellular constituent in said organism.

60. The method of claim 2 wherein said abundance value comprises an amount of phosphorylation of said cellular constituent.

61. The method of claim 1 wherein said one or more loci consist of at least two loci.

62. The method of claim 1, wherein said respective eQTL and said respective cQTL are deemd to be colocalized at a locus *Q* of said one or more loci when said respective eQTL

and said respective cQTL satisfy a pleiotropy test and, wherein failure of the pleiotropy test indicates that (i) the respective eQTL and the respective cQTL are two closely linked QTL, (ii) step (B) is not performed, and (iii) said cellular constituent is not determined to be causal for said trait of interest **T**.

63. The method of claim 62 wherein said pleiotropy test comprises comparing a model for a null hypothesis, indicating that said eQTL and said respective cQTL colocalize as a QTL, to a model for an alternative hypothesis, indicating that said eQTL and said respective cQTL are two closely linked QTL.

64. The method of claim 63 wherein said model for said null hypothesis is:

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} N + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}$$

wherein

N is a categorical random variable indicating the genotype at locus Q across said plurality of organisms;

$\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}$ is distributed as a bivariate normal random variable with mean $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and covariance matrix $\begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2 \\ \sigma_2\sigma_1 & \sigma_2^2 \end{pmatrix}$; and

μ_i and β_i are model parameters.

65. The method of claim 63 wherein said model for said alternative hypothesis is:

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} \beta_1 & \beta_2 \\ \beta_3 & \beta_4 \end{pmatrix} \begin{pmatrix} N_1 \\ N_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}$$

wherein

N_1 and N_2 are categorical random variables indicating the genotype at locus Q across said plurality of organisms;

$\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}$ is distributed as a bivariate normal random variable with mean $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and covariance matrix $\begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2 \\ \sigma_2\sigma_1 & \sigma_2^2 \end{pmatrix}$; and

μ_i and β_i are model parameters.

66. The method of claim 63 wherein said model for said alternative hypothesis is:

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} \beta_1 & \beta_2 \\ \beta_3 & \beta_4 \end{pmatrix} \begin{pmatrix} N_1 \\ N_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}$$

wherein

Q_1 and Q_2 are categorical random variables indicating the genotype at locus Q across said plurality of organisms;

$\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}$ is distributed as a bivariate normal random variable with mean $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and covariance matrix $\begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2 \\ \sigma_2\sigma_1 & \sigma_2^2 \end{pmatrix}$;

μ_i and β_i are model parameters; and one of the conditions (i) through (iv) is valid:

(i) $\beta_1 \neq 0$, $\beta_4 \neq 0$, $\beta_2 = 0$, and $\beta_3 = 0$;

(ii) $\beta_1 \neq 0$, $\beta_4 \neq 0$, $\beta_2 \neq 0$, and $\beta_3 = 0$;

(iii) $\beta_1 \neq 0$, $\beta_4 \neq 0$, $\beta_2 = 0$, and $\beta_3 \neq 0$; and

(iv) $\beta_1 \neq 0$, $\beta_4 \neq 0$, $\beta_2 \neq 0$, and $\beta_3 \neq 0$.

67. The method of claim 63 wherein said comparing comprises:

obtaining a first maximum likelihood estimate for the model for the null hypothesis by maximizing the loglikelihood for the model for the null hypothesis with respect to model parameters;

obtaining a second maximum likelihood estimate for the model for the alternative hypothesis by maximizing the loglikelihood for the model for the alternative hypothesis with respect to model parameters; and

forming a likelihood ratio test statistic between the first maximum likelihood estimate and said second maximum likelihood estimate to determine whether the model for the alternative hypothesis provides for a statistically significant better fit to the data than the model for the null hypothesis.

68. The method of claim 1 wherein said testing comprises considering a null test for causality having the relationship:

$$P(T,Q_{\bullet}|G) = P(T|G)P(Q_{\bullet}|G),$$

wherein

each function P is a probability density function;

T is a random variable for the trait of interest across said plurality of organisms;

$Q_{\bullet}$ is a genotype random variable for locus Q of said one or more loci across said plurality of organisms; and

G is said abundance pattern of said cellular constituent across said plurality of organisms.

69. The method of claim 68 wherein said testing comprises comparing said null test for causality, indicating that G is causal for T, to an alternative hypothesis, indicating that T and Q are dependent given G.

70. The method of claim 69 wherein said testing comprises optimizing the log likelihood ratio of said null hypothesis and said alternative hypothesis using maximum likelihood analysis.

71. The method of claim 1, the method further comprising repeating step (A) for each cellular constituent in a plurality of cellular constituents thereby identifying a candidate causative cellular constituent set, wherein each cellular constituent in said candidate causative cellular constituent set was identified in an instance of step (A) and wherein each cellular constituent in said plurality of cellular constituents that does not have a druggable domain is optionally excluded from said candidate causative cellular constituent set.

72. The method of claim 71 wherein a rank of a cellular constituent **i** in said candidate cellular constituent set is determined by an amount of genetic variation in the trait of interest T that is explained by the at least one eQTL of cellular constituent **i**.

73. The method of claim 71 wherein the amount of genetic variation in the trait of interest T that is explained by the at least one eQTL of cellular constituent **i** is determined by a joint analysis of the trait of interest at each one of the eQTL in said at least one eQTL.

74. The method of claim 1 wherein a determination that the cellular constituent is causal for the trait of interest **T** is validated by a gene knock-out experiment, a transgenic construction experiment, or an siRNA experiment.

75. The method of claim 1 wherein a further condidtion for finding that the cellular constituent is cuasal for said trait of interest **T** is that the variation in abundance levels of the cellular constituent across the plurality of organisms associates with the variation in said trait of interest **T** across said plurality of organisms.

76. The method of claim 75 wherein whether the association is present between (i) the variation in abundances level of the cellular constituent and (ii) the variation in said trait of interest **T** across the plurality of organisms is determined using a Pearson correlation, discriminant analysis or a regression model.

77. The method of claim 76 wherein a Pearson correlation is used and the association between (i) the variation in abundance levels of said cellular constituent and (ii) the variation in the trait of interest **T** across the plurality of organisms is found to be present when the Pearson correlation coefficient (p-value) is less than 0.00001.

78. The method of claim 76 wherein a Pearson correlation is used and the assocation between (i) the variation in abundance levels of said cellular constituent and (ii) the variation in the trait of interest **T** across the plurality of organisms is found to be present when the Pearson correlation coefficient (p-value) is less than 0.0001.

79. A method for determining whether a cellular constituent is causal for a trait of interest, the trait of interest **T** exhibited by one or more organisms in a plurality of organisms of a species, the method comprising:

(A) identifying one or more loci in the genome of said species, wherein each locus Q of said one or more loci is a site of colocalization for (i) a respective abundance quantitative trait locus (eQTL) genetically linked to a variation in abundance levels of the cellular constituent across the plurality of organisms and (ii) a respective clinical quantitative trait locus (cQTL) that is genetically linked to a variation in said trait of interest **T** across the plurality of organisms; and

(B) comparing, for one or more respective locus Q of said one or more loci, (i) a causative model, (ii) a reactive model and (iii) an independent model using a maximum likelihood approach, wherein

when, for each said compared locus Q of said one or more loci, the causative model gives rise to the largest likelihood relative to the corresponding reactive model and the corresponding independent model, said cellular constituent is deemed to be causal for said trait of interest.

80. The method of claim 79 wherein, for a given locus Q of said one or more loci, said causative model is defined as:

$$P(Q_{\bullet},G,T) = P(G|Q_{\bullet})P(T|G)$$

where $Q_{\bullet}$ is a genotype random variable for the locus Q across said plurality of organisms, G is said variation in abundance level of said cellular constituent across said plurality of organisms, and T is said variation of said trait of interest **T** across said plurality of organisms.

81. The method of claim 79 wherein, for a given locus Q of said one or more loci, said reactive model is defined as:

$$P(Q_{\bullet},G,T) = P(T|Q_{\bullet})P(G|T)$$

where $Q_{\bullet}$ a genotype random variable for the locus Q across said plurality of organisms, G is said variation in abundance level of said cellular constituent across said plurality of organisms, and T is a trait random variable for the trait of interest **T** across said plurality of organisms.

82. The method of claim 79, wherein, for a given locus Q of said one or more loci, said independent model is defined as:

$$P(Q_{\bullet},G,T) = P(T|Q_{\bullet})P(G|Q_{\bullet})$$

where $Q_{\bullet}$ is a geneotype random variable for the locus Q across said plurality of organisms, G is said variation in abundance level of said cellular constituent across said

plurality of organisms, and T is a trait random variable for the trait of interest **T** across said plurality of organisms.

83. The method of claim 79 wherein said maximum likelihood approach comprises maximizing said causative model, said reactive model, and said independent model using cellular constituent abundance data for said cellular constituent in said plurality of organisms, phenotype data for said trait of interest **T** in said plurality of organisms, and genotypic data at said locus Q in said plurality of organisms.

84. The method of claim 79 wherein

(i) a result of the maximum likelihood approach for said causative model for a given locus Q of said one or more loci Q is expressed in terms of a first Akaike Information Criterion;

(iii) a result of the maximum likelihood approach for said independent model for said given locus Q is expressed in terms of a second Akaike Information Criterion (AIC); and

(iii) a result of the maximum likelihood approach for said reactive model for said given locus Q is expressed in terms of a third Akaike Information Criterion; and wherein

the model associated with the lowest Akaike Information Criterion has the largest likelihood.

85. A computer program product for use in conjunction with a computer system, the computer program product comprising a computer readable storage medium and a computer program mechanism embedded therein, the computer program mechanism comprising:

a cQTL/eQTL overlap module that comprises instructions for identifying a cellular constituent **i** that has at least one abundance quantitative trait locus (eQTL) coincident with a respective clinical quantitative trait locus (cQTL) for a trait of interest at a respective locus Q; and

a causality test module that comprises instructions for testing, for one or more loci Q, whether (i) the genetic variation of said locus Q across all or a portion of a plurality of organisms of a species and (ii) the variation of the trait of interest across all or a portion of a plurality of organisms of said species are uncorrelated conditional on an abundance pattern of the cellular constituent **i** across the plurality of organisms.

86. A computer program product for use in conjunction with a computer system, the computer program product comprising a computer readable storage medium and a computer program mechanism embedded therein, the computer program mechanism comprising:

a cQTL/eQTL overlap module that comprises instructions for identifying a cellular constituent that has an abundance quantitative trait locus (eQTL) coincident with a clinical quantitative trait locus (cQTL) for a trait of interest at a respective loci Q wherein the trait of interest is exhibited by one or more organisms in a plurality of organisms of a species; and

a causality test module that comprises instructions for testing whether (i) a causative model, (ii) a reactive model or (iii) an independent model better describe the genetic relationship between the cellular constituent and the trait of interest, wherein

when the causative model gives rise to a better description, relative to the corresponding reactive model and the corresponding independent model, of the genetic relationship between the cellular constituent and the trait of interest, the cellular constituent is deemed to be causal for said trait of interest; and

a communication module for communicating the genetic relationship between the cellular constituent and the trait of interest.

87. A computer system comprising:

a central processing unit;

a memory, coupled to the central processing unit, the memory storing an cQTL/eQTL overlap module and a causality test module; wherein

the cQTL/eQTL overlap module comprises instructions for identifying a cellular constituent **i** that has at least one abundance quantitative trait locus (eQTL) coincident with a respective clinical quantitative trait locus (cQTL) for a trait of interest at a respective locus Q, wherein the trait of interest is exhibited by one or more organisms in a plurality of organisms of a species; and

the causality test module comprises instructions for testing, for one or more loci Q, whether (i) the genetic variation of said locus Q across all or a portion of a plurality of organisms of said species and (ii) the variation of the trait of interest across all or a portion of a plurality of organisms of said species are uncorrelated conditional on an abundance pattern of the cellular constituent **i** across the plurality of organisms.

88. A computer system comprising:

a central processing unit;

a memory, coupled to the central processing unit, the memory storing a cQTL/eQTL overlap module and a causality test module; wherein

the cQTL/eQTL overlap module comprises instructions for identifying a cellular constituent that has at least one abundance quantitative trait locus (eQTL) coincident with a respective clinical quantitative trait locus (cQTL) for a trait of interest at a respective locus Q in a plurality of loci, wherein the trait of interest is exhibited by a plurality of organisms of a species; and

the causality test module comprises instructions for testing, for one or more loci Q, whether (i) a causative model, (ii) a reactive model, or (iii) an independent model better describe the genetic relationship between the cellular constituent and the trait of interest, wherein,

when, the causative model gives rise to the largest likelihood relative to the corresponding reactive model and the corresponding independent model, said cellular constituent is deemed to be causal for said trait of interest; and

a communication module for communicating the genetic relationship between the cellular constituent and the trait of interest.

89. A method for determining whether a candidate molecule affects a body weight disorder associated with an organism, comprising:

(a) contacting a cell from said organism with, or recombinantly expressing within the cell from said organism, said candidate molecule;

(b) determining whether the RNA expression or protein expression in said cell of at least one open reading frame is changed in step (a) relative to the expression of said open reading frame in the absence of the candidate molecule, each said open reading frame being regulated by a promoter native to a nucleic acid sequence selected from the group consisting of SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 23 and homologs of each of the foregoing; and

(c) determining that the candidate molecule affects a body weight disorder associated with said organism when the RNA expression or protein expression of said at least one open reading frame is changed, or

determining that the candidate molecule does not affect a body weight disorder associated with said organism when the RNA expression or protein expression of said at least one open reading frame is unchanged.

90. The method of claim 89 wherein a cell from said organism contacted with the candidate molecule exhibits a lower expression level of a protein sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 10, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 22, SEQ ID NO: 24, and homologs of each of the forgoing, than a cell from said organism that is not contacted with said candidate molecule.

91. The method of claim 89, wherein step (b) comprises determining whether RNA expression is changed.

92. The method of claim 89, wherein step (b) comprises determining whether protein expression is changed.

93. The method of claim 89, wherein step (b) comprises determining whether RNA or protein expression of at least two of said open reading frames is changed.

94. The method of claim 89, wherein step (a) comprises contacting the cell with the candidate molecule, and wherein step (a) is carried out in a liquid high throughput-like assay.

95. The method of claim 89, wherein the cell comprises a promoter region of at least one gene selected from the group consisting of SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 23, and homologs of each of the foregoing, each promoter region being operably linked to a marker gene; and wherein step (b) comprises determining whether the RNA expression or protein expression of the marker gene(s) is changed in step (a) relative to the expression of said marker gene in the absence of the candidate molecule.

96. The method of claim 95, wherein the marker gene is selected from the group consisting of green fluorescent protein, red fluorescent protein, blue fluorescent protein,

luciferase, LEU2, LYS2, ADE2, TRP1, CAN1, CYH2, GUS, CUP1 and chloramphenicol acetyl transferase.

97. The method of claim 89, wherein said body weight disorder is obesity, anorexia nervosa, bulimia nervosa or cachexia.

98. A method of treating or preventing a body weight disorder comprising administering to a subject in which treatment is desired a therapeutically effective amount of a compound that antagonizes in the subject a protein comprising a sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 10, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 22, SEQ ID NO: 24 and homologs of each of the foregoing.

99. The method of claim 98 wherein said subject is human.

100. The method of claim 98 in which the compound:

(i) inhibits a function of one or more of the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 10, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 22, SEQ ID NO: 24, and homologs of each of the foregoing, and

(ii) is selected from the group consisting of:

an antibody that binds to one of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 10, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 22, SEQ ID NO: 24, and homologs of each of the foregoing or a fragment or derivative therefore containing the binding region thereof, or is selected from the group consisting of:

a nucleic acid complementary to the RNA produced by transcription of a gene encoding one of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 10, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 22, SEQ ID NO: 24, and homologs of each of the foregoing.

101. The method of claim 100 in which the compound that inhibits a function of one or more of the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 10, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO:

19, SEQ ID NO: 22, SEQ ID NO: 24, and homologs of each of the foregoing, is an oligonucleotide that:

(a) consists of at least six nucleotides;

(b) comprises a sequence complementary to at least a portion of an RNA transcript of a gene encoding one of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 10, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 22, SEQ ID NO: 24, and homologs of each of the foregoing; and

(c) is hybridizable to the RNA transcript under moderately stringent conditions.

102. A method of treating or preventing a body weight disorder comprising administering to a subject in which treatment is desired a therapeutically effective amount of a compound that enhances a function of one or more of the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 10, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 22, SEQ ID NO: 24, and homologs of each of the foregoing.

103. The method of claim 102 wherein said subject is human.

104. A method of diagnosing a disease or disorder or the predisposition to said disease or disorder, wherein the disease or disorder is characterized by an aberrant level of one of SEQ ID NO: 1 through SEQ ID NO: 24, or a homolog thereof, in a subject, the method comprising measuring the level of any one of SEQ ID NO: 1 through SEQ ID NO: 24, or a homolog thereof, in a sample derived from the subject, in which an increase or decrease in the level of one of SEQ ID NO: 1 through SEQ ID NO: 24, or a homolog thereof, in said sample, relative to the level of a corresponding one of said SEQ ID NO: 1 through SEQ ID NO: 24, or a homolog thereof, found in an analogous sample not having the disease or disorder, indicates the presence of the disease or disorder in the subject.

105. The method of claim 104 wherein the disease or disorder is a body weight disorder.

106. The method of claim 104 wherein the disease or disorder is obesity, anorexia nervosa, bulimia nervosa, or cachexia.

107. A method of diagnosing or screening for the presence of or predisposition for developing a disease or disorder involving a body weight disorder in a subject comprising detecting one or more mutations in at least one of SEQ ID NO: 1 through SEQ ID NO: 24, or a homolog thereof, in a sample derived from the subject, in which the presence of said one or more mutations indicates the presence of the disease or disorder or a predisposition for developing said disease or disorder.

108. A method for determining whether a first trait T_1 is causal for a second trait T_2 in a plurality of organisms of a species, the method comprising:

(A) identifying one or more loci in the genome of said species, wherein each locus Q of said one or more loci is a site of colocalization for (i) a respective quantitative trait locus (QTL_1) that is genetically linked to a variation in the first trait T_1 across the plurality of organisms and (ii) a respective quantitative trait locus (QTL_2) that is genetically linked to a variation in the second trait T_2 across said plurality of organisms; and

(B) testing, for each respective locus Q of said one or more loci, whether (i) a genetic variation Q_* of said respective locus Q across said plurality of organisms and (ii) said variation in said second trait T_2 across said plurality of organisms are correlated conditional on said variation in said first trait T_1 across said plurality of organisms,

wherein, when the genetic variation of (i) one or more loci Q tested in (B), and (ii) said variation in said second trait T_2 across said plurality of organisms are correlated conditional on said variation in said first trait T_1 across said plurality of organisms, said first trait T_1 is determined to be causal for said second trait T_2.

109. The method of claim 108, the method further comprising, prior to said identifying, a step of determining a respective QTL_1 at a locus Q of said one or more loci using a first quantitative trait locus (QTL) analysis, wherein said first QTL analysis uses a plurality of quantitative measurements of said first trait, and wherein each quantitative measurement in said plurality of quantitative measurements of said first trait is associated with an organism in said plurality of organisms.

110. The method of claim 109, the method further comprising a step of determining a respective QTL_2 at said locus Q using a second QTL analysis, wherein said second QTL analysis uses a plurality of quantitative measurements of said second trait, and wherein each quantitative measurement in said plurality of quantitative measurements of said second trait is associated with an organism in said plurality of organisms.

111. The method of claim 108, wherein said respective QTL_1 and said respective QTL_2 are deemed to be colocalized at a locus Q of said one or more loci when said respective QTL_1 and said respective QTL_2 are within 3 cM of the locus Q.

112. The method of claim 108, wherein said respective QTL_1 and said respective QTL_2 are deemed to be colocalized at a locus Q of said one or more loci when said respective QTL_1 and said respective QTL_2 are within 1 cM of the locus Q.

113. The method of claim 108 wherein said plurality of organisms is derived from a predetermined set of individuals.

114. The method of claim 108 wherein said plurality of organisms is derived from a predetermined set of strains.

115. The method of claim 114 wherein said set of strains is between 2 strains and 100 strains.

116. The method of claim 114 wherein said set of strains is between 5 strains and 500 strains.

117. The method of claim 114 wherein said set of strains is more than five strains.

118. The method of claim 114 wherein said set of strains is less than 1000 strains.

119. The method of claim 114 wherein said set of strains is diverse with respect to a complex phenotype associated with human disease.

120. The method of claim 114 wherein said set of strains is between 2 strains and 10 strains that, collectively, are diverse with respect to a complex phenotype associated with a human disease.

121. The method of claim 120 wherein said human disease is obesity, diabetes, atherosclerosis, metabolic syndrome, depression, anxiety, osteoporosis, bone development, asthma, or chronic obstructive pulmonary disease.

122. The method of claim 108 wherein said plurality of organisms is derived from crossing a predetermined set of strains.

123. The method of claim 122 wherein said plurality of organisms is an F_2 intercross, a backcross, or an F_2 random mating.

124. The method of claim 108 wherein the plurality of organisms is more than 1,000 organisms.

125. The method of claim 108 wherein the plurality of organism is between 100 organisms and 100,000 organisms.

126. The method of claim 108 wherein the plurality of organisms is less than 500,000 organisms.

127. The method of claim 108 wherein the plurality of organisms is between 5,000 and 25,000 organisms.

128. The method of claim 109, wherein

said first trait is abundance levels of a first cellular constituent and each quantitative measurement of said first trait is an abundance level of said first cellular constituent in an organism in said plurality of organisms; and

said second trait is abundance levels of a second cellular constituent and each quantitative measurement of said second trait is an abundance level of said second cellular constituent in an organism in said plurality of organisms.

129. The method of claim 128 wherein each said abundance level of said first cellular constituent is normalized and each said abundance level of said second cellular constituent is normalized

130. The method of claim 128 wherein

each said abundance level of said first cellular constituent is determined by measuring an amount of said first cellular constituent in one or more cells from an organism in said plurality of organisms; and

each said abundance level of said second cellular constituent is determined by measuring an amount of said second cellular constituent in one or more cells from an organism in said plurality of organisms.

131. The method of claim 128, wherein

each said amount of said first cellular constituent comprises an abundance of a first RNA in said one or more cells of said organism in said plurality of organisms; and

each said amount of said second cellular constituent comprises an abundance of a second RNA in said one or more cells of said organism in said plurality of organisms.

132. The method of claim 131, wherein

said abundance of said first RNA is measured by contacting a gene transcript array with said first RNA from said one or more cells of said organism, or with nucleic acid derived from said first RNA, wherein said gene transcript array comprises a positionally addressable surface with attached nucleic acids or nucleic acid mimics, wherein said nucleic acids or nucleic acid mimics are capable of hybridizing with said first RNA, or with nucleic acid derived from said first RNA; and.

said abundance of said second RNA is measured by contacting a gene transcript array with said second RNA from said one or more cells of said organism, or with nucleic acid derived from said second RNA, wherein said gene transcript array comprises a positionally addressable surface with attached nucleic acids or nucleic acid mimics, wherein said nucleic acids or nucleic acid mimics are capable of hybridizing with said second RNA, or with nucleic acid derived from said second RNA.

133. The method of claim 109, wherein said first QTL analysis comprises:

(i) testing for linkage between (a) the genotype of said plurality of organisms at a position in the genome of said species and (b) said plurality of quantitative measurements of said first trait;

(ii) advancing the position in said genome by an amount; and

(iii) repeating steps (i) and (ii) until all or a portion of the genome of said species has been tested.

134. The method of claim 110, wherein said second QTL analysis comprises:

(i) testing for linkage between (a) the genotype of said plurality of organisms at a position in the genome of said species and (b) said plurality of quantitative measurements of said second trait;

(ii) advancing the position in said genome by an amount; and

(iii) repeating steps (i) and (ii) until all or a portion of the genome of said species has been tested.

135. The method of claim 133 or 134, wherein said amount is less than 100 centiMorgans.

136. The method of claim 133 or 134, wherein said amount is less than 5 centiMorgans.

137. The method of claim 133 or 134, wherein said testing comprises performing linkage analysis or association analysis.

138. The method of claim 137, wherein said linkage analysis or association analysis generates a statistical score for said position in the genome of said species.

139. The method of claim 138, wherein said testing is linkage analysis and said statistical score is a logarithm of the odds (lod) score.

140. The method of claim 109, wherein said respective QTL_1 is represented by a lod score that is greater than 2.0.

141. The method of claim 110, wherein said respective QTL_2 is represented by a lod score that is greater than 2.0.

142. The method of claim 109, wherein said respective QTL_1 is represented by a lod score that is greater than 4.0.

143. The method of claim 110, wherein said respective QTL_2 is represented by a lod score that is greater than 4.0.

144. The method of claim 109 wherein each quantitative measurement in said plurality of quantitative measurements of said first trait is

an amount or a concentration of a first cellular constituent in one or more tissues of an organism in said plurality of organisms,

a cellular constituent activity level of said first cellular constituent in one or more tissues of an organism in said plurality of organisms, or

a state of cellular constituent modification of said first cellular constituent in one or more tissues of an organism in said plurality of organisms.

145. The method of claim 110 wherein each quantitative measurement in said plurality of quantitative measurements of said second trait is

an amount or a concentration of a second cellular constituent in one or more tissues of an organism in said plurality of organisms,

a cellular constituent activity level of said second cellular constituent in one or more tissues of an organism in said plurality of organisms, or

a state of cellular constituent modification of said second cellular constituent in one or more tissues of an organism in said plurality of organisms.

146. The method of claim 108, wherein said plurality of organisms is human.

147. The method of claim 109, wherein said respective QTL_1 and said respective QTL_2 are deemed to colocalize at a locus Q of said one or more loci when said respective QTL_1 and said respective QTL_2 are within 40 cM of the locus Q.

148. The method of claim 109, wherein said respective QTL_1 and said respective QTL_2 are deemed to colocalize at a locus Q of said one or more loci when said respective QTL_1 and said respective QTL_2 are within 10 cM of the locus Q.

149. The method of claim 108 wherein said one or more loci consist of at least two loci.

150. The method of claim 108, wherein said respective QTL_1 and said respective QTL_2 colocalize at a locus Q of said one or more loci when said respective QTL_1 and said respective QTL_2 satisfy a pleiotropy test and wherein failure of the pleiotropy test indicates that (i) the respective QTL_1 and the respective QTL_2 are two closely linked QTL, (ii) step (B) is not performed, and (iii) said first trait T_1 is not determined to be causal for said second trait T_2.

151. The method of claim 150 wherein said pleiotropy test comprises comparing a model for a null hypothesis, indicating that said respective QTL_1 and said respective QTL_2 colocalize as a QTL, to a model for an alternative hypothesis, indicating that said QTL_1 and said respective QTL_2 are two closely linked QTL.

152. The method of claim 151 wherein said model for said null hypothesis is:

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} N + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}$$

wherein

N is a categorical random variable indicating the genotype at locus Q across said plurality of organisms;

$\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}$ is distributed as a bivariate normal random variable with mean $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and covariance matrix $\begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2 \\ \sigma_2\sigma_1 & \sigma_2^2 \end{pmatrix}$; and

μ_i and β_i are model parameters.

153. The method of claim 151 wherein said model for said alternative hypothesis is:

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} \beta_1 & \beta_2 \\ \beta_3 & \beta_4 \end{pmatrix} \begin{pmatrix} N_1 \\ N_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}$$

wherein

N_1 and N_2 are categorical random variables indicating the genotype at locus Q across said plurality of organisms;

$\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}$ is distributed as a bivariate normal random variable with mean $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and covariance matrix $\begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2 \\ \sigma_2\sigma_1 & \sigma_2^2 \end{pmatrix}$; and

μ_i and β_i are model parameters.

155. The method of claim 152 wherein said model for said alternative hypothesis is:

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} \beta_1 & \beta_2 \\ \beta_3 & \beta_4 \end{pmatrix} \begin{pmatrix} N_1 \\ N_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}$$

wherein

Q_1 and Q_2 are categorical random variables indicating the genotype at locus Q across said plurality of organisms;

$\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}$ is distributed as a bivariate normal random variable with mean $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and covariance matrix $\begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2 \\ \sigma_2\sigma_1 & \sigma_2^2 \end{pmatrix}$;

μ_i and β_i are model parameters; and one of the conditions (i) through (iv) is valid:

(i) $\beta_1 \neq 0$, $\beta_4 \neq 0$, $\beta_2 = 0$, and $\beta_3 = 0$;

(ii) $\beta_1 \neq 0$, $\beta_4 \neq 0$, $\beta_2 \neq 0$, and $\beta_3 = 0$;

(iii) $\beta_1 \neq 0$, $\beta_4 \neq 0$, $\beta_2 = 0$, and $\beta_3 \neq 0$; and

(iv) $\beta_1 \neq 0$, $\beta_4 \neq 0$, $\beta_2 \neq 0$, and $\beta_3 \neq 0$.

155. The method of claim 151 wherein said comparing comprises:

obtaining a first maximum likelihood estimate for the model for the null hypothesis by maximizing the loglikelihood for the model for the null hypothesis with respect to model parameters;

obtaining a second maximum likelihood estimate for the model for the alternative hypothesis by maximizing the loglikelihood for the model for the alternative hypothesis with respect to model parameters; and

forming a likelihood ratio test statistic between the first maximum likelihood estimate and said second maximum likelihood estimate to determine whether the model for the alternative hypothesis provides for a statistically significant better fit to the data than the model for the null hypothesis.

156. The method of claim 108 wherein said testing comprises considering a null test for causality having the relationship:

$$P(T_2, Q_\bullet | T_1) = P(T_2 | G) P(Q_\bullet | T_1),$$

wherein

each function P is a probability density function;

T_2 is the variation of the second trait across said plurality of organisms;

$Q_\bullet$ is a genotype random variable for a locus Q of said one or more loci across said plurality of organisms; and

T_1 is the variation of the first trait across said plurality of organisms.

157. The method of claim 156 wherein said testing comprises comparing said null test for causality, indicating that said first trait T_1 is causal for said second trait T_2, to an alternative hypothesis, indicating that T_2 and Q are dependent given T_1.

158. The method of claim 157 wherein said testing comprises optimizing the log likelihood ratio of said null hypothesis and said alternative hypothesis using maximum likelihood analysis.

159. A computer program product for use in conjunction with a computer system, the computer program product comprising a computer readable storage medium and a computer program mechanism embedded therein, the computer program mechanism comprising:

a T_1/T_2 overlap module that comprises instructions for identifying one or more loci in the genome of a species, wherein each locus Q of said one or more loci is a site of colocalization for (i) a respective quantitative trait locus (QTL_1) that is genetically linked to a variation in a first trait T_1 across a plurality of organisms in said species and (ii) a respective quantitative trait locus (QTL_2) that is genetically linked to a variation in a second trait T_2 across said plurality of organisms; and

a causality test module that comprises instructions for testing, for one or more locus Q of said one or more loci, whether (i) a genotype random variable Q_* of the respective locus Q across the plurality of organisms and (ii) said variation in the second trait T_2 across the plurality of organisms are correlated conditional on the variation in said first trait T_1 across the plurality of organisms.

160. A computer system comprising:

a central processing unit;

a memory, coupled to the central processing unit, the memory storing an Q_1/Q_2 overlap module and a causality test module; wherein

the T_1/T_2 overlap module comprises instructions for identifying one or more loci in the genome of a species, wherein each locus Q of said one or more loci is a site of colocalization for (i) a respective quantitative trait locus (QTL_1) that is genetically linked to a variation in the first trait T_1 across a plurality of organisms of said species and (ii) a

respective quantitative trait locus (QTL_2) that is genetically linked to a variation in the second trait T_2 across said plurality of organisms; and

a causality test module that comprises instructions for testing, for one or more loci Q in the at least one locus, whether (i) a genotype random variable $Q_{\bullet}$ for the respective locus Q across the plurality of organisms and (ii) said variation in said second trait T_2 across said plurality of organisms are correlated conditional on the variation in the first trait T_1 across said plurality of organisms.

161. A method for determining whether a cellular constituent is causal for a trait of interest **T**, the trait of interest **T** exhibited by at least one organism in a plurality of organisms of a species, the method comprising:

(A) identifying a locus Q in the genome of said species that is a site of colocalization for (i) an abundance quantitative trait locus (eQTL) genetically linked to a variation in abundance levels of the cellular constituent across all or a portion of the plurality of organisms, and (ii) a clinical quantitative trait locus (cQTL) that is genetically linked to a variation in said trait of interest **T** across all or a portion of said plurality of organisms;

(B) quantifying a first coefficient of determination between (i) a variation in the clinical quantitative trait locus (cQTL) across all or a portion of the plurality of organisms, and (ii) a variation in the trait of interest **T** across all or a portion of said plurality of organisms; and

(C) quantifying a second coefficient of determination between (i) the variation in the clinical quantitative trait locus (cQTL) across all or a portion of the plurality of organisms, and (ii) the variation in the trait of interest **T** across all or a portion of said plurality of organisms, after conditioning on the variation of the abundance of the cellular constituent across all or a portion of said plurality of organisms; wherein

said cellular constituent is determined to be causal for said trait of interest **T** when said first coefficient of determination is other than zero and said second coefficient of determination cannot be distinguished from zero.

162. The method of claim 161 wherein said cellular constituent is determined to be causal for said trait of interest **T** when said first coefficient of determination is greater than a predetermined threshold amount.

163. The method of claim 162 wherein said predetermined threshold amount is 0.03.

164. The method of claim 162 wherein said predetermined threshold amount is 0.10.

165. The method of claim 161, wherein the eQTL is identified by a first quantitative trait locus (QTL) analysis, wherein said first QTL analysis uses a plurality of abundance statistics for said cellular constituent as a quantitative trait, and wherein each abundance statistic in said plurality of abundance statistics represents an abundance value for said cellular constituent in an organism in said plurality of organisms.

166. The method of claim 161, wherein the cQTL is identified by a second QTL analysis, wherein said second QTL analysis uses a plurality of phenotypic values, each phenotypic value in said plurality of phenotypic values corresponding to a quantitative measuremnet of the trait of interest **T** in an organism in said plurality of organisms.

167. The method of claim 161, wherein said eQTL and said cQTL are deemed to colocalize at said locus Q when said eQTL and said cQTL are within 3 cM of the locus Q.

168. The method of claim 161, wherein said eQTL and said cQTL are deemed to colocalize at said locus Q when said eQTL and said cQTL are within 1 cM of the locus Q.

169. The method of claim 161, wherein said first QTL analysis and said second QTL analysis each use a genetic map that represents the genome of said species.

170. The method of claim 169, the method further comprising, prior to said identifying, a step of constructing said genetic map from a set of genetic markers associated with said species.

171. The method of claim 170, wherein said set of genetic markers comprises single nucleotide polymorphisms (SNPs), microsatellite markers, restriction fragment length polymorphisms, short tandem repeats, DNA methylation markers, sequence length polymorphisms, random amplified polymorphic DNA, amplified fragment length polymorphisms, or simple sequence repeats.

172. The method of claim 171, wherein genotype data are used in said constructing and wherein said genotype data comprise knowledge of which alleles, for each marker in said set of genetic markers, are present in each organism in said plurality of organisms.

173. The method of claim 161 wherein the plurality of organisms is between 100 organisms and 100,000 organisms.

174. The method of claim 161 wherein the plurality of organisms is less than 500,000 organisms.

175. The method of claim 161 wherein the plurality of organisms is between 5,000 and 25,000 organisms.

176. The method of claim 165, wherein said first QTL analysis comprises:

(i) testing for linkage between (a) a genotype of all or a portion of said plurality of organisms at a position in the genome of said species and (b) said plurality of abundance statistics for said cellular constituent;

(ii) advancing the position in said genome by an amount; and

(iii) repeating steps (i) and (ii) until all or a portion of the genome of said species has been tested.

177. The method of claim 176, wherein said amount is less than 100 centiMorgans.

178. The method of claim 176, wherein said amount is less than 5 centiMorgans.

179. The method of claim 176, wherein said testing comprises performing linkage analysis or association analysis.

180. The method of claim 176, wherein said linkage analysis or association analysis generates a statistical score for said position in the genome of said species.

181. The method of claim 180, wherein said testing is linkage analysis and said statistical score is a logarithm of the odds (lod) score.

182. The method of claim 181, wherein said respective eQTL is represented by a lod score that is greater than 2.0.

183. The method of claim 181, wherein said respective eQTL is represented by a lod score that is greater than 4.0.

184. The method of claim 166, wherein said second QTL analysis comprises:

(i) testing for linkage between (a) a genotype of said plurality of organisms at a position in the genome of said species and (b) said plurality of phenotypic values;

(ii) advancing the position in said genome by an amount; and

(iii) repeating steps (i) and (ii) until all or a portion of the genome of said species has been tested.

185. The method of claim 184, wherein said amount is less than 100 centiMorgans.

186. The method of claim 184, wherein said amount is less than 5 centiMorgans.

187. The method of claim 184, wherein said testing comprises performing linkage analysis or association analysis.

188. The method of claim 187, wherein said linkage analysis or association analysis generates a statistical score for said position in the genome of said species.

189. The method of claim 188, wherein said testing is linkage analysis and said statistical score is a logarithm of the odds (lod) score.

190. The method of claim 189, wherein said respective cQTL is represented by a lod score that is greater than 2.0.

191. The method of claim 189, wherein said respective cQTL is represented by a lod score that is greater than 4.0.

192. The method of claim 161, wherein said plurality of organisms is human.

193. The method of claim 161, wherein said trait of interest **T** is a complex trait.

194. The method of claim 193, wherein said complex trait is characterized by an allele that exhibits incomplete penetrance in said species.

195. The method of claim 193, wherein said complex trait is a disease that is contracted by said at least one organism in said plurality of organisms, and wherein said organism inherits no predisposing allele to said disease.

196. The method of claim 193, wherein said complex trait arises when one or more of a plurality of different genes in the genome of said species is mutated.

197. The method of claim 193, wherein said complex trait requires the simultaneous presence of mutations in a plurality of genes in the genome of said species.

198. The method of claim 193, wherein said complex trait is a phenotype that does not exhibit Mendelian recessive or dominant inheritance attributable to a single gene locus.

199. The method of claim 193, wherein said complex trait is asthma, ataxia telangiectasia, bipolar disorder, cancer, common late-onset Alzheimer's disease, diabetes, heart disease, hereditary early-onset Alzheimer's disease, hereditary nonpolyposis colon cancer, hypertension, infection, maturity-onset diabetes of the young, mellitus, migraine, nonalcoholic fatty liver, nonalcoholic steatohepatitis, non-insulin-dependent diabetes mellitus, obesity, polycystic kidney disease, psoriases, schizophrenia, or xeroderma pigmentosum.

200. The method of claim 161, wherein said eQTL and said cQTL are deemed to colocalize at a locus Q of said one or more loci when said eQTL and said cQTL are within 40 cM of the locus Q.

201. The method of claim 161, wherein said eQTL and said cQTL are deemed to colocalize at a locus Q of said one or more loci when said eQTL and said cQTL are within 10 cM of the locus Q.

202. The method of claim 165 wherein each said abundance value comprises an amount of said cellular constituent in a tissue of an organism in said plurality of organisms, a

concentration of said cellular constituent in a tissue of an organism in said plurality of organisms, a cellular constituent activity level for said cellular constituent in a tissue of an organism in said plurality of organsims, or a state of modification of said cellular constituent in an organism in said plurality of organisms.

203. The method of claim 165 wherein each said abundance value comprises a degree of phosphorylation of said cellular constituent in a tissue of an organism in said plurality of organisms.

204. The method of claim 161, wherein said eQTL and said cQTL are deemed to colocalize at said locus Q when said eQTL and said cQTL satisfy a pleiotropy test, and wherein failure of the pleiotropy test indicates that the eQTL and the cQTL are two closely linked QTL and said cellular constituent is not determined to be causal for said trait of interest **T**.

205. The method of claim 204 wherein said pleiotropy test comprises comparing a model for a null hypothesis, indicating that said eQTL and said cQTL colocalize as a QTL, to a model for an alternative hypothesis, indicating that said eQTL and said respective cQTL are two closely linked QTL.

206. The method of claim 205 wherein said model for said null hypothesis is:

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} N + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}$$

wherein

N is a categorical random variable indicating the genotype at locus Q across said plurality of organisms;

$\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}$ is distributed as a bivariate normal random variable with mean $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and covariance matrix $\begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2 \\ \sigma_2\sigma_1 & \sigma_2^2 \end{pmatrix}$; and

μ_i and β_i are model parameters.

207. The method of claim 205 wherein said model for said alternative hypothesis is:

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} \beta_1 & \beta_2 \\ \beta_3 & \beta_4 \end{pmatrix} \begin{pmatrix} N_1 \\ N_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}$$

wherein

N_1 and N_2 are categorical random variables indicating the genotype at locus Q across said plurality of organisms;

$\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}$ is distributed as a bivariate normal random variable with mean $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and covariance matrix $\begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2 \\ \sigma_2\sigma_1 & \sigma_2^2 \end{pmatrix}$; and

μ_i and β_i are model parameters.

208. The method of claim 205 wherein said model for said alternative hypothesis is:

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} \beta_1 & \beta_2 \\ \beta_3 & \beta_4 \end{pmatrix} \begin{pmatrix} N_1 \\ N_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}$$

wherein

Q_1 and Q_2 are categorical random variables indicating the genotype at locus Q across said plurality of organisms;

$\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}$ is distributed as a bivariate normal random variable with mean $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and covariance matrix $\begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2 \\ \sigma_2\sigma_1 & \sigma_2^2 \end{pmatrix}$;

μ_i and β_i are model parameters; and one of the conditions (i) through (iv) is valid:

(i) $\beta_1 \neq 0$, $\beta_4 \neq 0$, $\beta_2 = 0$, and $\beta_3 = 0$;

(ii) $\beta_1 \neq 0$, $\beta_4 \neq 0$, $\beta_2 \neq 0$, and $\beta_3 = 0$;

(iii) $\beta_1 \neq 0$, $\beta_4 \neq 0$, $\beta_2 = 0$, and $\beta_3 \neq 0$; and

(iv) $\beta_1 \neq 0$, $\beta_4 \neq 0$, $\beta_2 \neq 0$, and $\beta_3 \neq 0$.

209. The method of claim 205 wherein said comparing comprises:

obtaining a first maximum likelihood estimate for the model for the null hypothesis by maximizing the loglikelihood for the model for the null hypothesis with respect to model parameters;

obtaining a second maximum likelihood estimate for the model for the alternative hypothesis by maximizing the loglikelihood for the model for the alternative hypothesis with respect to model parameters; and

forming a likelihood ratio test statistic between the first maximum likelihood estimate and said second maximum likelihood estimate to determine whether the model for the alternative hypothesis provides for a statistically significant better fit to the data than the model for the null hypothesis.

210. The method of claim 161 wherein a determination that the cellular constituent is causal for the trait of interest **T** is validated by a gene knock-out experiment, a transgenic construction experiment, or an siRNA experiment.

211. A method for determining whether a first trait T_1 is causal for a second trait T_2 in a plurality of organisms of a species, the method comprising:

(A) identifying a locus Q in the genome of said species that is a site of colocalization for (i) a quantitative trait locus (QTL_1) that is genetically linked to a variation in the first trait T_1 across all or a portion of the plurality of organisms and (ii) a quantitative trait locus (QTL_2) that is genetically linked to a variation in the second trait T_2 across all or a portion of said plurality of organisms;

(B) quantifying a first coefficient of determination between (i) a genetic variation Q_* of said locus Q across all or a portion of said plurality of organisms and (ii) said variation in said first trait T_1 across all or a portion of said plurality of organisms; and

(C) quantifying a second coefficient of determination between (i) said genetic variation Q^* of said locus Q across all or a portion of said plurality of organsisms and (ii) said variation in said first trait T_1 across all or a portion of said plurality of organsims, after conditioning on said variation in said second trait T_2 across all or a portion of said plurality of organisms, wherein

said first trait T_1 is deemed to be causal for said second trait T_2 when said first coefficient of determination is other than zero and said second coefficient of determination cannot be distinguished from zero.

212. The method of claim 211 wherein said cellular constituent is deemed to be causal for said trait of interest **T** when said first coefficient of determination is greater than a predetermined threshold amount.

213. The method of claim 212 wherein said predetermined threshold amount is 0.03.

214. The method of claim 212 wherein said predetermined threshold amount is 0.10.

215. The method of claim 211, wherein said QTL_1 and said QTL_2 are deemed to colocalize at said locus Q when said QTL_1 and said QTL_2 are within 3 cM of the locus Q.

216. The method of claim 211, wherein said QTL_1 and said QTL_2 are deemed to colocalize at said locus Q when said QTL_1 and said QTL_2 are within 1 cM of the locus Q.

217. The method of claim 211 wherein the plurality of organisms is between 100 organisms and 100,000 organisms.

218. The method of claim 211 wherein the plurality of organisms is less than 500,000 organisms.

219. The method of claim 211 wherein the plurality of organisms is between 5,000 and 25,000 organisms.

220. The method of claim 211 wherein said plurality of organisms is human.

221. The method of claim 211, wherein said first trait T_1 is a complex trait.

222. The method of claim 221, wherein said complex trait is characterized by an allele that exhibits incomplete penetrance in said species.

223. The method of claim 221, wherein said complex trait is a disease that is contracted by said at least one organism in said plurality of organisms, and wherein said organism inherits no predisposing allele to said disease.

224. The method of claim 221, wherein said complex trait arises when one or more of a plurality of different genes in the genome of said species is mutated.

225. The method of claim 221, wherein said complex trait requires the simultaneous presence of mutations in a plurality of genes in the genome of said species.

226. The method of claim 221, wherein said complex trait is a phenotype that does not exhibit Mendelian recessive or dominant inheritance attributable to a single gene locus.

227. The method of claim 221 wherein said complex trait is asthma, ataxia telangiectasia, bipolar disorder, cancer, common late-onset Alzheimer's disease, diabetes, heart disease, hereditary early-onset Alzheimer's disease, hereditary nonpolyposis colon cancer, hypertension, infection, maturity-onset diabetes of the young, mellitus, migraine, nonalcoholic fatty liver, nonalcoholic steatohepatitis, non-insulin-dependent diabetes mellitus, obesity, polycystic kidney disease, psoriases, schizophrenia, or xeroderma pigmentosum.

228. The method of claim 211 wherein said QTL_1 and said QTL_2 are deemed to colocalize at a locus Q of said one or more loci when said QTL_1 and said QTL_2 are within 40 cM of the locus Q.

229. The method of claim 211 wherein said QTL_1 and said QTL_2 are deemed to colocalize at a locus Q of said one or more loci when said QTL_1 and said QTL_2 are within 10 cM of the locus Q.

230. The method of claim 211 wherein said QTL_1 and said QTL_2 are deemed to colocalize at said locus Q when said QTL_1 and said QTL_2 satisfy a pleiotropy test and wherein failure of the pleiotropy test indicates that the QTL_1 and the QTL_2 are two closely linked QTL and said first trait T_1 is not determined to be causal for said second trait T_2.

231. The method of claim 230 wherein said pleiotropy test comprises comparing a model for a null hypothesis, indicating that said QTL_1 and said QTL_2 colocalize as a QTL, to a model for an alternative hypothesis, indicating that said QTL_1 and said QTL_2 are two closely linked QTL.

232. The method of claim 231 wherein said model for said null hypothesis is:

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} N + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}$$

wherein

N is a categorical random variable indicating the genotype at locus Q across said plurality of organisms;

$\begin{pmatrix}\varepsilon_1\\ \varepsilon_2\end{pmatrix}$ is distributed as a bivariate normal random variable with mean $\begin{pmatrix}0\\0\end{pmatrix}$ and covariance matrix $\begin{pmatrix}\sigma_1^2 & \sigma_1\sigma_2\\ \sigma_2\sigma_1 & \sigma_2^2\end{pmatrix}$; and

μ_i and β_i are model parameters.

233. The method of claim 231 wherein said model for said alternative hypothesis is:

$$\begin{pmatrix}y_1\\ y_2\end{pmatrix}=\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}+\begin{pmatrix}\beta_1 & \beta_2\\ \beta_3 & \beta_4\end{pmatrix}\begin{pmatrix}N_1\\ N_2\end{pmatrix}+\begin{pmatrix}\varepsilon_1\\ \varepsilon_2\end{pmatrix}$$

wherein

N_1 and N_2 are categorical random variables indicating the genotype at locus Q across said plurality of organisms;

$\begin{pmatrix}\varepsilon_1\\ \varepsilon_2\end{pmatrix}$ is distributed as a bivariate normal random variable with mean $\begin{pmatrix}0\\0\end{pmatrix}$ and covariance matrix $\begin{pmatrix}\sigma_1^2 & \sigma_1\sigma_2\\ \sigma_2\sigma_1 & \sigma_2^2\end{pmatrix}$; and

μ_i and β_i are model parameters.

234. The method of claim 231 wherein said model for said alternative hypothesis is:

$$\begin{pmatrix}y_1\\ y_2\end{pmatrix}=\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}+\begin{pmatrix}\beta_1 & \beta_2\\ \beta_3 & \beta_4\end{pmatrix}\begin{pmatrix}N_1\\ N_2\end{pmatrix}+\begin{pmatrix}\varepsilon_1\\ \varepsilon_2\end{pmatrix}$$

wherein

Q_1 and Q_2 are categorical random variables indicating the genotype at locus Q across said plurality of organisms;

$\begin{pmatrix}\varepsilon_1\\ \varepsilon_2\end{pmatrix}$ is distributed as a bivariate normal random variable with mean $\begin{pmatrix}0\\0\end{pmatrix}$ and covariance matrix $\begin{pmatrix}\sigma_1^2 & \sigma_1\sigma_2\\ \sigma_2\sigma_1 & \sigma_2^2\end{pmatrix}$;

μ_i and β_i are model parameters; and one of the conditions (i) through (iv) is valid:

(i) $\beta_1 \neq 0$, $\beta_4 \neq 0$, $\beta_2 = 0$, and $\beta_3 = 0$;

(ii) $\beta_1 \neq 0$, $\beta_4 \neq 0$, $\beta_2 \neq 0$, and $\beta_3 = 0$;

(iii) $\beta_1 \neq 0$, $\beta_4 \neq 0$, $\beta_2 = 0$, and $\beta_3 \neq 0$; and

(iv) $\beta_1 \neq 0$, $\beta_4 \neq 0$, $\beta_2 \neq 0$, and $\beta_3 \neq 0$.

235. The method of claim 231 wherein said comparing comprises:

obtaining a first maximum likelihood estimate for the model for the null hypothesis by maximizing the loglikelihood for the model for the null hypothesis with respect to model parameters;

obtaining a second maximum likelihood estimate for the model for the alternative hypothesis by maximizing the loglikelihood for the model for the alternative hypothesis with respect to model parameters; and

forming a likelihood ratio test statistic between the first maximum likelihood estimate and said second maximum likelihood estimate to determine whether the model for the alternative hypothesis provides for a statistically significant better fit to the data than the model for the null hypothesis.

236. A method for identifying a quantitative trait locus for a trait that is exhibited by a plurality of organisms in a population, comprising:

(a) dividing said population into a plurality of sub-populations using a classification scheme that classifies each organism in said population into at least one of said subpopulations, wherein said classification scheme is derived from a plurality of cellular constituent measurements for each of a plurality of respective cellular constituents that are obtained from each said organism and wherien said classification scheme uses a classifier constructed using boosting or adaptive boosting; and

(b) for at least one sub-population in said plurality of sub-populations, performing quantitative genetic analysis on said sub-population in order to identify said quantitative trait locus for said trait.

237. The method of claim 236, wherein said cellular constituent measurements from each said organism are transcriptional state measurements or translational state measurements.

238. The method of claim 237, wherein said translational state measurements are performed using an antibody array or two-dimensional gel electrophoresis.

239. The method of claim 236, wherein said respective plurality of cellular constituents comprises a plurality of metabolites and said plurality of cellular constituent measurements are derived by a cellular phenotypic technique.

240. The method of claim 239, wherein said cellular phenotypic technique comprises a metabolomic technique wherein a plurality of levels of metabolites in each said organism is measured.

241. The method of claim 240, wherein said metabolites comprise an amino acid, a metal, a soluble sugar, or a complex carbohydrate.

242. The method of claim 240, wherein said plurality of levels of metabolites is measured by use of pyrolysis mass spectrometry, fourier-transform infrared spectrometry, Raman spectrometry, gas chromatography-mass spectroscopy, capillary electrophoresis, high pressure liquid chromatography / mass spectroscopy (HPLC/MS), liquid chromatography (LC)-electrospray mass spectroscopy, or cap-LC-tandem electrospray mass spectroscopy.

243. The method of claim 236 wherein said plurality of cellular constituent measurements comprise gene expression levels, abundance of mRNA, protein expression levels, or metabolite levels.

244. The method of claim 236, wherein said trait is characterized by an allele that exhibits incomplete penetrance in said population.

245. The method of claim 236, wherein said trait is a disease that is contracted by an organism in said population, and wherein said organism inherits no predisposing allele to said disease.

246. The method of claim 236, wherein said trait arises when any of a plurality of different genes in the genome of said plurality of organisms is mutated.

247. The method of claim 236, wherein said trait requires the simultaneous presence of mutations in a plurality of genes in the genome of said plurality of organisms.

248. The method of claim 236, wherein said trait is associated with a high frequency of disease-causing alleles in said population.

249. The method of claim 236, wherein said trait is a phenotype that does not exhibit Mendelian recessive or dominant inheritance attributable to a single gene locus.

250. The method of claim 236, wherein said trait is asthma, ataxia telangiectasia, bipolar disorder, cancer, common late-onset Alzheimer's disease, diabetes, heart disease, hereditary early-onset Alzheimer's disease, hereditary nonpolyposis colon cancer, hypertension, infection, maturity-onset diabetes of the young, mellitus, migraine, nonalcoholic fatty liver, nonalcoholic steatohepatitis, non-insulin-dependent diabetes mellitus, obesity, polycystic kidney disease, psoriases, schizophrenia, or xeroderma pigmentosum.

251. The method of claim 236, wherein said plurality of cellular constituent measurements from each said organism comprises the measurement of the cellular constituent levels of ten or more cellular constituents in each said organism.

252. The method of claim 236, wherein said plurality of cellular constituent measurements from each said organism comprises the measurement of the cellular constituent levels of one thousand or more cellular constituent levels in each said organism.

253. The method of claim 236, wherein said dividing further comprises verifying the division of said population into said plurality of sub-populations.

254. The method of claim 236, wherein said quantitative genetic analysis is performed using a method selected from the group consisting of a linkage analysis, a quantitative trait locus (QTL) analysis method that uses said plurality of cellular constituent measurements as a phenotypic trait, and an association analysis.

255. The method of claim 254, wherein said quantitative genetic analysis is performed using said QTL analysis, said QTL analysis method comprising:

(a) clustering QTL data from a plurality of QTL analyses to form a QTL interaction map, wherein

each QTL analysis in said plurality of QTL analyses is performed for a gene **G** in a plurality of genes in the genome of said plurality of organisms using a genetic marker map and a quantitative trait in order to produce said QTL data, wherein, for each QTL analysis, said quantitative trait comprises an expression statistic for the gene **G**, for which the QTL analysis has been performed, for each organism in said plurality of organisms; and wherein

said genetic marker map is constructed from a set of genetic markers associated with said plurality of organisms; and

(b) analyzing said QTL interaction map to identify said QTL associated with said quantitative trait.

256. The method of claim 255, which further comprises, prior to said clustering step, a step of constructing said genetic marker map from said set of genetic markers associated with said plurality of organisms.

257. The method of claim 255, which further comprises, prior to said clustering step, a step of performing each said QTL analysis in said plurality of QTL analyses.

258. The method of claim 255, wherein said expression statistic for said gene **G** is computed by a method comprising transforming an expression level measurement of said gene **G** from each organism in said plurality of organisms.

259. The method of claim 258, wherein said step of transforming an expression level measurement of said gene **G** comprises normalizing the expression level measurement of said gene **G** in order to form said expression statistic.

260. The method of claim 259, wherein normalizing the expression level measurement of said gene **G** in order to form said expression statistic is performed by a normalization technique selected from the group consisting of Z-score of intensity, median intensity, log median intensity, Z-score standard deviation log of intensity, Z-score mean absolute deviation of log intensity, calibration DNA gene set, user normalization gene set, ratio median intensity correction, and intensity background correction.

261. The method of claim 255, wherein each said QTL analysis comprises:

(i) testing for linkage between a position in the genome of said plurality of organisms, and the quantitative trait used in the QTL analysis;

(ii) advancing the position in said genome by an amount; and

(iii) repeating steps (i) and (ii) until all or a portion of the genome has been tested.

262. The method of claim 261, wherein said amount is less than 100 centiMorgans.

263. The method of claim 261, wherein said QTL data produced from each respective QTL analysis comprises a statistical score computed at each said position.

264. The method of claim 261, the method further comprising creating a QTL vector for each quantitative trait tested in said chromosome, wherein said QTL vector comprises a statistical score for each position tested by the QTL analysis corresponding to the quantitative trait.

265. The method of claim 264, wherein said clustering of QTL data comprises clustering each said QTL vector.

266. The method of claim 264, wherein a similarity metric that is used as a basis for said clustering is a Euclidean distance, a squared Euclidean distance, a Euclidean sum of squares, a Manhattan metric, a Pearson correlation coefficient, or a squared Pearson correlation coefficient, and wherein the similarity metric is computed between QTL vector pairs.

267. The method of claim 261 or 265, wherein said clustering of QTL data comprises applying a hierarchical clustering technique, applying a k-means technique, applying a fuzzy k-means technique, applying a Jarvis-Patrick clustering, applying a self-organizing map technique, or applying a neural network technique.

268. The method of claim 267, wherein said clustering of QTL data comprises applying a hierarchical clustering technique, wherein said hierarchical clustering technique is an agglomerative clustering procedure.

269. The method of claim 268, wherein said agglomerative clustering procedure is a nearest-neighbor algorithm, a farthest-neighbor algorithm, an average linkage algorithm, a centroid algorithm, or a sum-of-squares algorithm.

270. The method of claim 267, wherein said hierarchical clustering technique is a divisive clustering procedure.

271. The method of claim 261, wherein said step of analyzing said QTL interaction map comprises filtering the QTL interaction map in order to obtain a candidate pathway group.

272. The method of claim 271, wherein said filtering in order to obtain said candidate pathway group comprises selecting those QTL for said candidate pathway group that interact most strongly with another QTL in said QTL interaction map.

273. The method of claim 272, wherein said QTL that interact most strongly with another QTL in said QTL interaction map are those QTL in said QTL interaction map that share a correlation coefficient with another QTL in said quantitative trait locus interaction map that is higher than 75% of all correlation coefficients computed between QTL in said quantitative trait locus interaction map.

274. The method of claim 272, the method further comprising fitting a multivariate statistical model to said candidate pathway group in order to test the degree to which each QTL making up the candidate pathway group belongs in the candidate pathway group.

275. The method of claim 274, wherein said multivariate statistical model simultaneously considers multiple quantitative traits.

276. The method of claim 274, wherein said multivariate statistical model looks for epistatic interactions between QTL in said candidate pathway group.

277. The method of claim 261, wherein said set of genetic markers comprises a single nucleotide polymorphism (SNP), a microsatellite marker, a restriction fragment length polymorphism, a short tandem repeat, a DNA methylation marker, or a sequence length polymorphism.

278. The method of claim 261, wherein pedigree data is used in step (b) of claim 236, and wherein said pedigree data shows one or more relationships between organisms in said plurality of organisms.

279. The method of claim 236, wherein said plurality of organisms is human.

280. The method of claim 236, wherein said dividing step (a) comprises:

(i) partitioning said population into a plurality of phenotypic groups using phenotypic data for all or a portion of said plurality of organisms;

(ii) identifying a set of extreme organisms in said plurality of phenotypic groups that represent a phenotypic extreme;

(iii) identifying cellular constituents within said plurality of cellular constituents, wherein each respective identified cellular constituent has the property that cellular constituent measurements for the respective cellular constituent obtained from said set of extreme organisms discriminate all or a portion of said plurality of phenotypic groups;

(iv) constructing a classifier using a probability distribution derived from all or a portion of said identified cellular constituents and a boosting technique or an adaptive boosting technique.

281. The method of claim 280 wherein said phenotypic data comprises a binary event.

282. The method of claim 280 wherein said phenotypic data comprises more than one phenotypic measurement for each organism in said population.

283. The method of claim 280 wherein said phenotypic data comprises a determination as to whether each organism in said plurality of organisms exhibits a trait, and said partitioning step (i) comprises placing an organism in said plurality of organisms in a first phenotypic group when said organism exhibits said trait and placing an organism in said plurality of organisms in a second phenotypic group when said organism does not exhibit said trait.

284. The method of claim 280 wherein an organism represents said phenotypic extreme when it is the top 30^{th} or bottom 30^{th} percentile of said population with respect to a phenotype exhibited by said population.

285. The method of claim 280 wherein an organism represents said phenotypic extreme when it is the top 10^{th} or bottom 10^{th} percentile of said population with respect to a phenotype exhibited by said population.

286. The method of claim 280 wherein said set of extreme organisms is more than 5 organisms.

287. The method of claim 280 wherein said set of extreme organisms is between 2 and 100 organisms.

288. The method of claim 280 wherein said set of extreme organisms is less than 1000 organisms.

289. The method of claim 280 wherein said identifying step (iii) comprises subjecting a plurality of cellular constituent measurements for a predetermined cellular constituent to a t-test, wherein said plurality of cellular constituent measurements is obtained from said set of extreme organisms.

290. The method of claim 280 wherein said identifying step (iii) comprises subjecting a group of identified cellular constituents within said plurality of cellular constituents to multivariate analysis.

291. The method of claim 280 wherein said cellular constituents identified in step (iii) are reduced prior to said constructing step (iv).

292. The method of claim 291 wherein said cellular constituents identified in step (iii) are reduced by stepwise regression, all-possible-subset regression, principal component analysis, or multiple-discriminant analysis.

293. The method of claim 291 wherein said cellular constituents identified in step (iii) are reduced by a stochastic search method.

294. The method of claim 293 wherein said stochastic search method is simulated annealing or a genetic algorithm.

295. A computer program product for use in conjunction with a computer system, the computer program product comprising a computer readable storage medium and a computer program mechanism embedded therein, the computer program mechanism comprising:

a classification module for dividing a plurality of organisms in a population into a plurality of sub-populations using a classification scheme that classifies each organism in said population into at least one of said sub-populations, wherein said classification scheme is derived from a plurality of cellular constituent measurements for each of a plurality of respective cellular constituents that are obtained from each said organism in said population and wherein said classification scheme uses a classifier constructed using boosting or adaptive boosting;

a quantitative genetic analysis module that, for at least one sub-population in said plurality of sub-populations, performs quantitative genetic analysis on said sub-population in order to identify a quantitative trait locus for a complex trait that is exhibited by one or more organisms in said plurality of organisms.

296. A computer system for identifying a quantitative trait locus for a complex trait that is exhibited by a plurality of organisms in a population, the computer system comprising:

a central processing unit;

a memory, coupled to the central processing unit, the memory storing a classification module and a quantitative genetic analysis module; wherein

the classification module includes instructions for dividing a plurality of organisms in a population into a plurality of sub-populations using a classification scheme that classifies each organism in said population into at least one of said sub-populations, wherein said classification scheme is derived from a plurality of cellular constituents measurements for each of a plurality of respective cellular constituents that are obtained from each said organism in said population and wherein said classification scheme uses a classifier constructed using boosting or adaptive boosting; and

the quantitative genetic analysis module includes instructions that, for at least one sub-population in said plurality of sub-populations, performs quantitative genetic analysis on said sub-population in order to identify said quantitative trait locus for said complex trait.